UDC 57.088.1

MASS SPECTROMETRY-BASED APPROACHES TO CHARACTERIZE *MYCOBACTERIUM TUBERCULOSIS*

Atavliyeva S., Tarlykov P., Zholdybayeva E., Ramankulov Ye.

National Center for Biotechnology
13/5, Korgalzhyn road, Astana, 010000, Kazakhstan
atavlievas@gmail.com

ABSTRACT

Mass spectrometry has greatly contributed to the study and understanding of the pathogenesis of human tuberculosis. Current methods of mass spectrometry have been rapidly evolving over the past two decades in response to the limitations of early proteomic studies. Modern proteomic research includes protein secretion, activation, degradation, and modifications, since post-translational modifications are an additional step in the evolution of pathogenic mycobacteria that gain virulence. The other application of mass spectrometry is in epidemiological studies of mycobacteria, in particular, determining the genetic spoligotype based on matrix-assisted laser desorption/ionization. This review explores *M. tuberculosis* studies carried out by mass spectrometry, including proteomic profiling with the shotgun-based approach and targeted proteomics. In addition, the most important post-translational modifications studied by mass spectrometry are described.

Keywords: mass spectrometry, *Mycobacterium tuberculosis*, proteome, spoligotyping**.**

INTRODUCTION

The use of mass spectrometry (MS) in microbial studies is becoming an indispensable tool, covering a wide range of topics in basic and applied research. Bacteria are ideal model organisms to develop a system biology approach that combines data from several areas, including genomics, transcriptomics, proteomics, and metabolomics. Recent proteomics studies carried out by the means of mass spectrometry has been successful in experiments on host-pathogen interaction and mixed microbial communities [1].

The strategy used by mycobacteria to invoke the pathological process in the human body involves many mechanisms; therefore, systems biology approach is needed. The complexity of the pathological process in combination with the evolving drug resistance of *M. tuberculosis* contributes to the status of tuberculosis as the deadliest infectious disease. According to the World Health Organization report (2016), the causative agent of tuberculosis has resulted in 1.3 million deaths worldwide [2].

In the study of *M. tuberculosis*, most system-level studies have been focused on the genomics of the pathogen. The complete genomic sequence of the laboratory strain *Mycobacterium tuberculosis* H37Rv was deciphered in 1998. The study identified about 4,000 open reading frames pointing to 16% of the predicted proteins that were specific for mycobacteria, most of which were hypothetically identified as virulence and pathogenicity factors [3]. Subsequently, the predicted virulence factors were studied using mutagenesis to identify mycobacterial genes necessary for survival *in vivo* by the method of comparative genomics and hybridization of transposon sites [4]. The progress achieved as a result of the

genetic analysis of mycobacteria has deepened the current understanding of the molecular basis of *M. tuberculosis*, but the genomic data of the microorganism could not provide the answers on mycobacterial environmental plasticity and nature of its antibiotic resistance. Proteomics, in turn, seeks to provide information on the dynamic regulation of the bacteria in changing environmental conditions, allowing a detailed comparison of proteomes, including protein interactions, modifications and cell localization [5, 6].

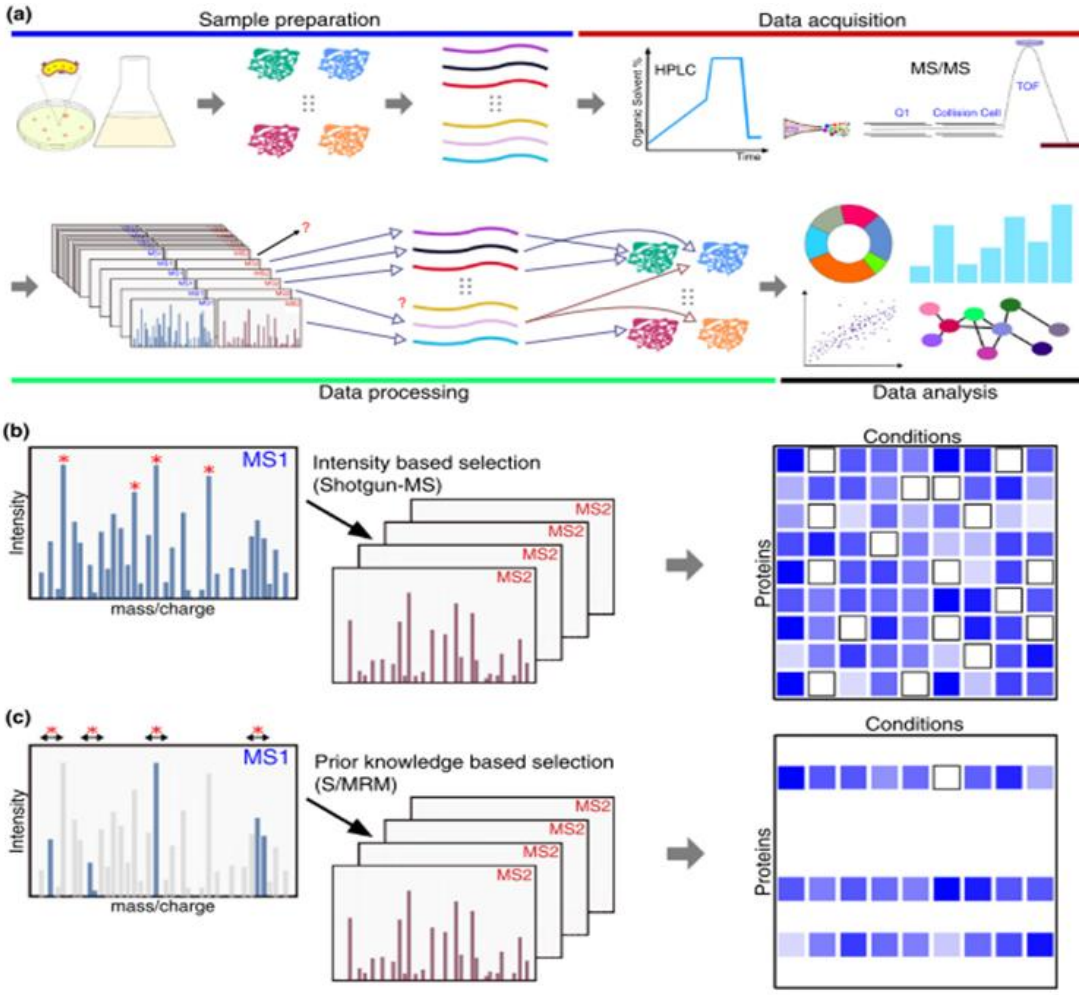
Proteomics-based mass spectrometry makes it possible to identify and quantitatively analyze complex protein mixtures, which constitute proteome. Currently, proteomic profiling methods are based on liquid chromatography with tandem mass spectrometry (LC-MS/MS) and are divided into top-down and bottom-up approaches. The basis of downstream proteomics is the formation of intact proteins, which are fragmented in a mass spectrometer for broader analysis. In turn, bottom-up proteomics is more common and relies on proteolytic cleavage of proteins to peptides, followed by processing for MS identification or quantitative analysis [7].

One of the goals of the functional proteomics is post-translational modification (PTM) profiling, the study of signaling pathways and protein complexes that is responsible for bacterial response to environmental adaptation, production of virulence and pathogenicity factors. Protein modifications contribute to the ability of mycobacteria to rapidly change the activity of proteins. The most important modifications known to date are glycosylation, phosphorylation, and acetylation [8].

Proteomic profiling

Microorganisms have always been in the foreground of proteomics. For instance, the very first use of top-down proteomics approach has been applied to prokaryotes [1]. Proteomic profiling had started with the use of two-dimensional gel electrophoresis (2D-GE) followed by identification by MS. The first studies of the *M. tuberculosis* proteome using 2D-GE made it possible to determine approximately 100 proteins. In 2004, a complex analysis of the proteome of *M. tuberculosis* H37Rv was performed by Schmidt et al. As a result, 108 proteins were found in the lysate [9]. Due to the limited resolution of the 2D-GE method, novel proteomic methods based on MS were developed for the simultaneous analysis of up to several thousand proteins.

The use of high-performance liquid chromatography with tandem MS (LC-MS/MS) increased the number of identified proteins to 3000, and from this point on the bottom-up proteomics has been developed [8]. Bottom-up proteomics uses two basic approaches of LC-MS/MS: shotgun proteomics and targeted proteomics (fig. 1).



(a) - The workflow shows the four major steps of a bottom-up proteomic study consisting of sample preparation, data acquisition, processing and analysis. The significant difference between various proteomic methods is the data acquisition step and associated data analysis illustrated in panel b–d. (b) - In Shotgun-MS ionized peptides are isolated for MS2 scan based on their intensity. The scheme depicts the selected precursors using asterisk signs. The final data matrix has a considerable amount of missing values. (c) - S/MRM are knowledge based data acquisition methods where one has to obtain chromatographic and mass spectrometric coordinates of peptides of interest in order to define the isolation scheme. As these methods do not rely on MS1 scans for isolating precursors, they are considered as data independent acquisition methods.

Fig. 1. Schematic overview on bottom-up proteomics and its three major modes [6].

In both approaches, proteins are enzymatically cleaved to peptides; the resulting mixture of peptides is separated using reverse phase chromatography, which is directly connected to electrospray ionization. After the initial determination of the mass-to-charge ratio of molecular ions (MS1), peptide ions are transferred to a mass spectrometer vacuum, where they undergo fragmentation in the gas phase, the MS/MS spectra (MS2) contain information for identifying the amino acid sequence and quantifying specific peptides [7]. Fragmentation of proteins into peptides is usually carried out by trypsin; however, chymotrypsin is better suited for transmembrane proteins of mycobacteria. It has been shown that chymotrypsin increases protein coverage, which is explained by the fact that transmembrane proteins have hydrophobic sequences, which makes them inaccessible for trypsinolysis [10].

Shotgun proteomics

In proteomics, a mass spectrometer operates in data-dependent acquisition mode (DDA), where the most intense ions detected during MS1 scanning are selected for fragmentation and MS2 analysis. Shotgun proteomics is aimed at reaching the full proteome, it allows to identify

and analyze 1000 proteins in one run. The DDA mode allows running unknown proteins. Moreover, all analyzed proteins can be identified in one run [7].

About ten years ago, mass spectrometry has contributed to new ways to study tuberculosis, therefore, most of the data on proteomic profiling accumulated during this time are formed using this method. Thus, MS facilitated comparative proteomic profiling between avirulent (H37Ra) and virulent (H37Rv) strains of mycobacteria, including membrane proteins. About 1,700 proteins were identified, of which 1,300 were common to both strains, 278 were present exclusively in the proteome of the H37Rv strain, and 193 proteins were identified only in the *M. tuberculosis* H37Ra strain. Among the proteins whose concentration was higher in the virulent H37Rv strain, the SecF protein involved in intercellular metabolism and three proteins of the ABC carrier family were identified, indicating a possible role in the virulence of the strains of transport channels in the cell membrane of the bacillus [11].

Previous studies of the comparative genomics of the H37Ra and H37Rv strains identified significant differences in the PE/PPE/PE-PGRS family of genes, including SNVs (single nucleotide variant) and indels. Proteins of the PE-/PPE family, named for the structural features of the N-terminus of the Pro-Glu (PE) and Pro-Pro-Glu (PPE) motifs, are acidic proteins with a high content of glycine and alanine and are characterized by repetitive sequences. Therefore, these proteins are not prone to tryptic cleavage, however, 24-fold overexpression and 4-fold hyperphosphorylation of PE31 protein in the log phase of H37Rv compared with H37Ra was determined. The PE15 protein, which is involved in evading the immune response of the human body and promotes the survival of the bacillus *in vivo*, was found to be more expressed in the logarithmic phase than in the stationary phase of the H37Rv strain. Thus, transmembrane proteins play an important role in the pathogenicity of *M. tuberculosis* [10, 12].

In a similar manner, Mattow et al. compared the membrane proteome of *M. tuberculosis* H37Rv and the attenuated strain of *M. bovis* BCG vaccine (bacillus Calmette-Guérin). It is known that the genomes of the two strains are 99.9% identical. The comparative genomics analysis identified deletions in the BCG genome, called regions of difference (RD). The absence of RDs in the BCG genome makes it non-pathogenic. In turn, proteomic profiling identified a number of proteins corresponding to these regions in *M. bovis* that may contribute to the virulence of H37Rv compared to BCG. Among certain distinct sites, two virulence factors were observed, namely, ESAT-6 and CFP-10 [13]. Interestingly, the ESAT-6 protein is found only in H37Rv strains, but not in H37Ra, and mutations have been identified in genes of ESAT-6-like proteins in the H37Ra strain. The relationship of ESAT-6 and CFP-10 proteins is determined through an antigen (Esat-6) forming a heterodimer complex with the latter; there is an assumption that the ESAT-6/CFP-10 is associated with the virulence of *M. tuberculosis* and interferes with phagosome and lysosome fusion within the host macrophages, thereby preventing the destruction of mycobacterial cells [14].

In a recent study, Gunawardena et al. compared the membrane proteomic profiles of H37Rv and BCG, where several proteins associated with RD genomic regions in MTB and BCG were identified. Additionally, a decrease in SigK protein (rv0445) in BCG, which is a sigma factor (initiation factors that promote the attachment of RNA polymerase to specific initiation sites), was determined. The decrease in secretion is explained by a point mutation in the BCG start codon ATA that makes MTB start codon ATG. The difference in codons reduces the efficiency of translation in *M. bovis* [15]. Thus, proteomics contributed to the determination of the proteomic basis of the virulence of mycobacteria when comparing pathogenic and attenuated strains.

Targeted proteomics

Target proteomics removes the limitations imposed by shotgun proteomics, such as reproducibility of results and quantitative analysis of given sets of proteins. Thus, the device operates in the directional mode.

The targeted proteomics method depends on the collection mode of the mass spectra: selected reaction monitoring (SRM), as well as multiple reaction monitoring (MRM). A mass spectrometer selects the required precursor ion from a mixture of peptides by m/z values, then the selected peptide is fragmented and measured quantitatively by the intensity of fragment ions [2, 16].

Due to the high selectivity and sensitivity of the methods, S/MRM allows obtaining accurate and reproducible measurement data of proteins covering a large dynamic range. The dynamic range of expression of *M. tuberculosis* proteins is estimated to be more than four orders of magnitude with the majority of proteins covering only two orders of magnitude, which is significantly less than the dynamic range of human plasma. Modern mass spectrometers have the ability to detect and quantify peptides spanning from four to five orders of magnitude [7, 8].

An important requirement of the S/MRM method is the knowledge of the mass spectra of the target proteins, peptides, prior to their quantitative evaluation. For each protein of interest, it is necessary to determine by experiment several pairs of fragments (precursor ions) with the quantitative data. For each precursor ion, it is necessary to create high-quality fragmentation spectra, which is complicated if the protein is present in low amounts. Currently, this task can be solved more easily by downloading the results to the Atlas SRM public database. Today, SRM Atlas contains data from targeted human proteomic analyzes, *M. tuberculosis*, and *S. cerevisiae*, with a 97% mycobacterial proteome coverage [17].

The SRM Atlas database was used in a study by Schubert et al. to measure the response to hypoxia in the survival regulator (dormancy survival regulon (DosR)), which is involved in the persistence and resistance of mycobacteria. DosR contains 53 proteins, of which 85% were identified in the hypoxic state quantitatively using SRM. The HspX and TB31.7 proteins were overexpressed; in addition, nine DosR regulon proteins were identified that were previously below the MS detection limit and could not be detected by the shotgun method [16]. In turn, Picotti et al. also used SRM analysis and identified proteins in *S. cerevisiae* that had previously avoided LC-MS/MS detection [18].

Methods of targeted proteomics are effective when applied to the verification and validation of data from comparative research proteomics or to obtain absolute quantitative information of selected proteins that are of interest to the researcher.

The study of *M. tuberculosis* by methods of proteomic profiling contributes to the reliable identification and quantification of proteins, which is important for understanding the physiology and pathology of tuberculosis. As described above, proteomics identifies thousands of proteins in a single instrument run, but the reproducibility and quantitative accuracy of the method is reduced. Targeted proteomics tracks quantitative information about a predetermined set of proteins, directing the mass spectrometer to select peptides, which are the proteins studied. In general, the developed methods of proteomics complement each other, since each method is suitable for a specific research purpose.

Analysis of post-translation modifications in *M. tuberculosis*

Post-translational modifications (PTM) are covalent modifications of proteins that occur on the amino acid side chains or N- or C-termini. PTMs lead to the formation of proteoforms, representing various functional states of the protein. PTMs carry information about cell regulation and cell response to changes in the environment, controlling numerous processes in the cell. The study of PTMs using MS has been successfully applied in research since a modification in protein leads to characteristic mass shifts. Thus, identification is based on the specific increase in the mass of the modified amino acid residue or the presence of a characteristic spectrum of the ionic fragment obtained as a result of fragmentation [6, 7, 8].

Mass spectrometry is the only method that is able to evaluate complex PTM profile. Before the appearance of MS, PTM was studied by isolating a single modified protein or were identified by chance. When studying PTM in mycobacteria, genes containing homologous eukaryotic information containing serine/threonine protein kinases and prokaryotic ubiquitin-like protein

were originally found [19]. Modified proteins are usually produced in the cell in a low copy number and are not identified by the method of shotgun proteomics since specific enrichment is required before MS analysis [5, 19].

Eight types of protein modifications were found in the mycobacterial proteome (O-glycosylation, phosphorylation, methylation, acetylation, lipidation, deamidation, N-formylation, ubiquitination), we will focus on the three most significant: phosphorylation, glycation, acetylation. It has been shown that at least 30% of *M. tuberculosis* proteins undergo one or more types of PTM [6].

Phosphorylation

Phosphorylation is PTM, which is involved in the regulation of biochemical pathways, altering the function of the protein associated with pathogenicity. It is a reversible modification, which is catalyzed by protein kinases and removed by protein phosphatases. When protein is phosphorylated, phosphate is added to the reactive groups of the amino acid side chain; hydroxyl amino acids are exposed: serine, threonine, and tyrosine (Ser/Thr/Tyr), as well as histidine and aspartate. However, due to the acidic instability of histidine and aspartate phosphorylation, these amino acids are not currently available for MS-based proteomics [8, 20].

Various Ser/Thr/Tyr kinases are found in prokaryotes. *Mycobacterium tuberculosis* is able to survive in the host macrophages for decades, evading the immune response. In order to survive in adverse macrophage conditions, mycobacteria use effective regulatory mechanisms, such as Ser/Thr/Tyr phosphorylation [5, 20].

The phosphoproteomics by MS includes an enrichment step with since phosphorylated peptides are underrepresented when compared to the whole proteome. Enrichment for phosphopeptides is usually achieved by a combination of strong cation exchange and TiO_2 chromatography or by immunoprecipitation with specific antibodies prior to MS analysis. When analyzing LC-MS/MS, phosphate groups remain attached to the peptide and are available for identification [5, 8].

When the phosphorylation level of the proteins in *M. smegmatis* and *M. bovis* was compared, an increased level of phosphorylation was observed in the BCG strain in comparison with *M. smegmatis*. The authors concluded that phosphorylation patterns are species-specific and occur on different peptides [21]. As a result of the study of the hypervirulent Beijing *M. tuberculosis* strain, 414 phosphorylated sites were identified on 214 proteins, of which 252 sites were not previously identified in the H37Rv strain [20].

For the first time, a large-scale study of the *M. tuberculosis* H37Rv phosphoproteome was carried out by Prisic et al. Mycobacteria were cultured under conditions of oxidative stress, which is similar to macrophage conditions during an immune response. 152 samples were analyzed using MS, 506 phosphorylation events were identified on 301 proteins. The ratio (40:60) of serine and threonine phosphorylation, respectively, was determined. Phosphorylated tyrosine sites were not detected [22].

Eleven serine/threonine protein kinases (STPKs) are encoded in the *M. tuberculosis* genome. Nine of which are transmembrane kinases, called PknA-L. The surface of the mycobacterium cell is the first barrier of contact with the human body, which reinforces the argument that phosphorylation of membrane proteins increases the virulence of the bacteria, and inhibition of kinase activity is used as a basis for a therapeutic approach.

It is known that the secretion of the PnkG gene promotes the survival of mycobacteria in macrophages by inhibiting phagosome-lysosomal fusion, after being absorbed by macrophages [8]. PknH protein kinase phosphorylates the DosR regulator, which emphasizes the importance of kinases during the transition to hypoxia-induced rest in the macrophage. In turn, the chemical inhibition of PknB prevents the growth of *Mycobacterium* after the DosR state. Thus, the quantitative determination of the phosphorylation status of Ser/Thr proteins and protein kinases is an indicator of the strain virulence [5]. Tyrosine kinases were discovered in the bacterial proteome and were found to be important in the regulation of homeostasis. Tyrosine

phosphorylation in mycobacteria was first reported by Kusebauch et al. Thirty tyrosine sites on 17 mycobacterial proteins were identified with the Ser:Thr:Tyr ratio (34:62:4) [23].

Cell signaling is directly related to the number of phosphorylation events, therefore, an increasing number of phosphorylated sites may play a role in the virulence of different *M. tuberculosis* strains.

Acetylation

Acetylation is a process of addition of acetyl group to the lysine side chain, to Ser/Thr residues or to the N-terminus using acetyl coenzyme A (Ac-Coa). Acetylation is a reversible PTM. Proteins with acetylated lysine residues may exhibit changes in protein stability, interaction, localization, and function. It is known that acetylation in prokaryotes determines pathogenicity [24]. Like many other post-translational modifications, acetylation causes a shift in 2D-GE, which was used to isolate modified proteins. Gel-based separation in combination with MS has shown that acetylation of ESAT-6 leads to differential binding to CFP-10 [25]. This earlier study has suggested searching for acetylation sites mediating the virulent properties of mycobacteria.

Acetylation was also studied in *M. tuberculosis* H37Ra using the LC-MS/MS. Acetylated peptides were enriched using specific antibodies, and identification in MS was determined by the mass shift of the modified amino acid. In the study of acetylome of *M. tuberculosis* H37Ra strain, 226 acetylation sites in 137 proteins were identified. Functional analysis of these proteins has shown that many acetylated proteins are indeed involved in fatty acid metabolism, which implies the potential role of acetylation in determining the virulence of *M. tuberculosis* [26].

Interestingly, a recent study has indicated that lysine acetylation plays an important role in carbon metabolism. Mycobacteria encounter phagolysosomes in macrophages and survive in conditions with low nutrient content, an acidic environment (pH 5.8-6.2) and hypoxia. One of the mechanisms of its sustainability lies in altering the central carbon metabolism and using the fatty acids of the host cell as a substrate instead of carbohydrates. The metabolic pathway of the transition of fatty acids to carbon contains isocitrate-lyase and malate synthase enzymes. An MS-based large-scale review on acetylation of the *M. tuberculosis* virulence strain revealed important additional virulence factors. The isocitrate-lyase (ICL1) enzyme is involved in central carbon metabolism and is a key pathogenicity factor. The enzyme has three identified acetylated lysine residues (K322, K331, and K392). The acetylation of K392 has been shown to increase the activity of the ICL1 enzyme, whereas the acetylation of K322 reduces its activity. Therefore, lysine acetylation plays a critical role in the regulation of the mechanism of ICL1 [27].

Another mechanism of adaptation to hypoxia and the transition of the mycobacterium to the latent state are also regulated by acetylation. Acetylome was quantitatively analyzed to compare the acetylation profile during aeration and hypoxia and changes were detected in 377 sites in 269 mycobacterial proteins. The deacetylation of DosR in position K182 was of particular interest. It has contributed to the hypoxia reaction in *M. tuberculosis* and enhanced the transcription of genes targeting DosR. According to the study, hypoxia causes DosR deacetylation, which, in turn, increases its ability to bind to DNA and cause transcription of target genes, allowing the mycobacteria to go into a latent state [28].

The role of acetylation in adapting the virulent strain of *M. tuberculosis* to various conditions may be of interest for the development of drugs intended for latent mycobacteria.

Glycosylation

Enzymatic attachment of sugar residues to proteins leads to post-translational modification called glycosylation. Carbohydrates are usually present on cell membranes, where they mediate specific interactions between cells. The study of glycosylated mycobacterial proteins contributes to the identification of virulence factors since the glycoproteins of bacterial cell membranes are used by *M. tuberculosis* to penetrate the cells of the human body [5].

O-mannosylation of the MPT32 antigen (38 kDa), which interacts and infects the host macrophages, was found in tuberculosis bacilli. O-mannosylation is catalyzed by membrane-associated proteins O-mannosyl transferases, which transfer mannose residues from lipid carriers to proteins. The study of *M. bovis* cell surface lipoproteins using MS facilitated the detection of MPB83 protein with mannosylated sites on Thr48 and Thr49 [29]. Using LC-MS/MS Gonzalez-Zamorano et al. identified 41 glycoproteins, of which 31 were previously unknown. All proteins were highly different in function. In a similar work, Smith et al. investigated *M. tuberculosis* cultures and found that O-mannosylation affects 13 extracytoplasmic proteins. Therefore, O-mannosylation may be an important modification of secreted proteins acting as virulent factors [30]. Indeed, *in vivo* experiments in a mouse model showed that the mannosyltransferase gene has inactivated a decrease in the pathogenicity of *M. tuberculosis* [31].

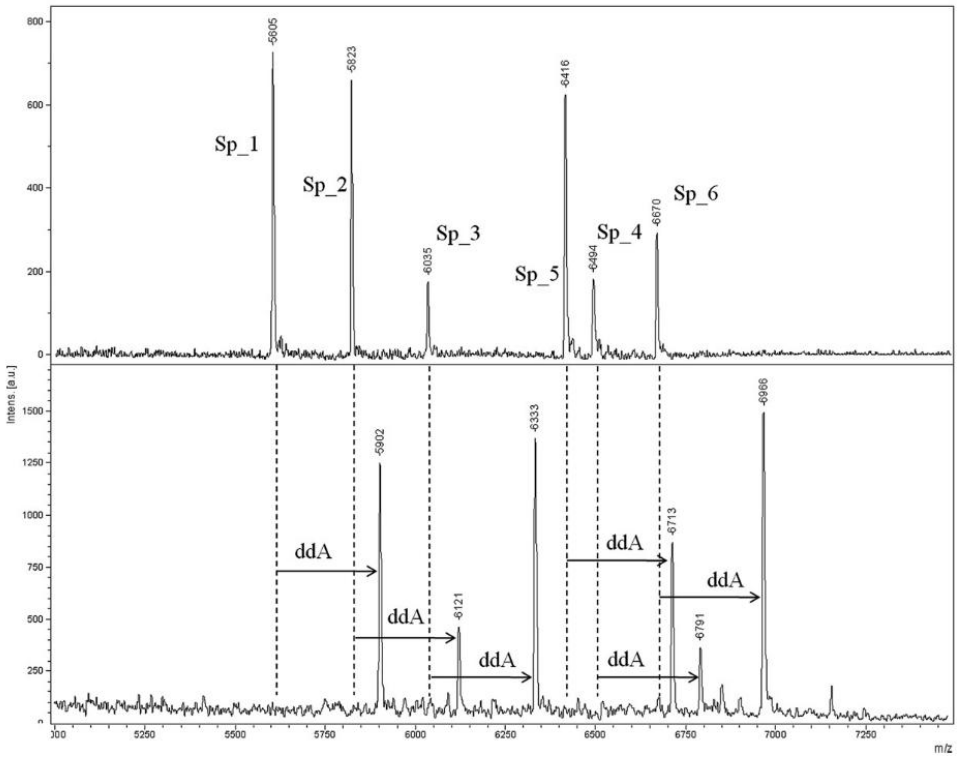
To date, the role of PTMs of mycobacteria is not well understood. The ability to detect PTM by mass spectrometry will be gradually improving with new technology. Therefore, the interpretation of the obtained data and comparison with the previously studied modifications is extremely important when focusing on clinical use.

Spoligotyping on MALDI-TOF MS

Spoligotyping is a method of molecular characterization of *Mycobacterium* strains of the tuberculosis complex (MTBC), which allows you to quickly get the results of typing and identification for epidemiological purposes. The analysis is based on a chromosomal locus containing a varying number of identical 36-bp direct repeats (DR), alternating with unique sequences (spacers) from 30-bp to 40-bp in length. Strains differ in the number of DRs and the presence or absence of 43 individual spacers [32].

The standard method of spoligotyping is based on the identification of 43 spacer sequences by amplifying the DR region and subsequent hybridization of PCR products with nylon membranes on which synthetic oligonucleotide probes specific for each spacer are immobilized. Next, the membranes undergo chemiluminescent staining, where hybridization of PCR products with probes is determined by spots [33].

Matrix-assisted laser desorption/ionization using a time-of-flight mass analyzer (MALDI-TOF MS) is an analytical tool for measuring m/z of various biomolecules (proteins, lipids, nucleic acids). It was successfully used to detect the spoligotype. This analysis includes the target amplification of the DR locus followed by extension PCR with primer and one of the dideoxynucleotides (ddNTP). Oligonucleotides are designed and multiplexed to avoid cross-hybridization. The method was designed in a way that oligonucleotides were extended and terminated with ddNTP and their possible termination product has an expected mass in the spectrum. The presence of a spacer in a sample leads to an elongation of the oligonucleotide by one base and, accordingly, a change in the signal in the MS, unlike the absence of a spacer in this region. In the absence of a spacer, a signal of the mass of the non-elongated oligonucleotide will be the only one visible mass on the spectrum (fig. 2) [33, 34].



Probes Sp_1, Sp_2, Sp_4 b Sp_5 are forward directed, probes Sp_3 and Sp_6 are reverse. Mass spectra of original probes (upper part) and products of reaction (lower part) are displayed. The picture presents the shift of all original probes by ddA, which shows the presence of spacers in the analyzed locus.

Fig. 2. Schematic diagram of chain extension reaction on the example of spoligotype spacer 1-6 detecting [34].

When analyzing the spoligotyping of 326 samples based on MALDI-TOF MS, a match of 96.6% (315/326) was found, as noted by the authors, most of the discordant samples were identified as a result of ambiguous interpretation of weak hybridization signals in blot analysis [33].

A while later, the possibility of spoligotyping for the determination of mycobacterial strains based on MALDI-TOF MS was confirmed on 383 clinical samples. Of these, 348 were MTBC strains and 35 strains were nontuberculous mycobacteria, represented by 16 different species. The dominant family was found to be Beijing (74.9%), then, in descending order LAM9 (8.2%), T1 (3.5%), Haarlem3 (2.6%), T4 (2.4%) and T1_RUS2 (1.7%). The results of the study based on MS completely coincided with the data obtained using commercial spoligotyping kit [34].

The use of MALDI-TOF MS for spoligotype detection is an alternative, fast and reproducible method for identifying and typing MTBC.

CONCLUSION

Today, the mass spectrometry-based study of *M. tuberculosis* allowed in-depth proteome characterization of clinical isolates. Moreover, accurate mass spectrometry-based profiling of the biological molecules has found application in the analysis DNA (spoligotyping), targeted proteomics that tracks quantitative information about a predetermined set of proteins, and analysis of PTMs. The growing area of *M. tuberculosis* proteomics has the potential to advance our understanding of strain variation in circulating *M. tuberculosis* strains.

Methods of proteomic profiling in combination with PTM studies contributed to the determination of the virulence factors and pathogenicity of mycobacteria. Thus, the use of MS in the study of *M. tuberculosis* contributes to the determination of candidates for therapeutic intervention. One of the requirements for successful treatment of tuberculosis is the rapid

identification of mycobacteria. The method of spoligotyping for identification is in demand for epidemiological, diagnostic purposes and was tested in an appendix to the MS. The results show that MALDI-TOF MS has the potential of being a fast and reproducible platform for identifying and typing MTBC. This could provide novel insights into biological mechanisms and pathogenesis of various *M. tuberculosis* isolates.

Acknowledgments

This study was supported by the Ministry of Education and Science of the Republic of Kazakhstan (grant No. AP05130238).

REFERENCES

1. Otto A., Becher D., Schmidt F. Quantitative proteomics in the field of microbiology. *Proteomics*, 2014, vol. 14, no. 4-5, pp. 547-565. doi:10.1002/pmic.201300403

2. World Health Organization. Global tuberculosis report 2017. Available at: http://www.who.int/tb/publications/global_report/en/ (accessed at 9 September 2018).

3. Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C. et al. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature*, 1998, vol. 393, pp. 537–544. doi: 10.1038/31159

4. Sassetti C., Rubin E.J. Genetic requirements for mycobacterial survival during infection. *Proc. Natl. Acad. Sci. USA*, 2004, vol. 100, pp. 12989–12994.

5. Calder B., Soares N.C., de Kock E., Blackburn J.M. Mycobacterial proteomics: analysis of expressed proteomes and post-translational modifications to identify candidate virulence factors. *Expert Rev Proteomics*, 2015, vol. 12, no. 1, pp. 21-35. doi:10.1586/14789450.2015.1007046

6. Banaei-Esfahani A., Nicod C., Aebersold R., Collins B.C. Systems proteomics approaches to study bacterial pathogens: application to *Mycobacterium tuberculosis*. *Curr Opin Microbiol.*, 2017, vol. 39, pp. 64-72. doi:10.1016/j.mib.2017.09.013

7. Aebersold R., Mann M. Mass-spectrometric exploration of proteome structure and function. *Nature*, 2016; vol. 537, pp. 347-355. doi:10.1038/nature19949

8. Kaufmann S.H.E., Aebersold R., Schubert O.T., Mouritsen J., Gengenbacher M. *Mycobacterium tuberculosis* in the Proteomics Era. *Microbiol Spectr*, 2014, vol. 2, pp. 1-19. doi:10.1128/microbiolspec.MGM2-0020-2013

9. Schmidt F., Donahoe S., Hagens K., Mattow J., Schaible U. E., Kaufmann S. H., Aebersold R., Jungblut P. R. Complementary analysis of the *Mycobacterium tuberculosis* proteome by two-dimensional electrophoresis and isotope-coded affinity tag technology. *Mol. Cell. Proteomics*, 2004, vol. 3, no. 1, pp. 24–42.

10. Bell C., Smith G.T., Sweredoski M.J., Hess S. Characterization of the *Mycobacterium tuberculosis* proteome by liquid chromatography mass spectrometry-based proteomics techniques: A comprehensive resource for tuberculosis research. *J Proteome Res.*, 2012, vol. 11, no. 1, pp. 119-130. doi:10.1021/pr2007939

11. Malen H., De Souza G., Pathak S, Softeland T., Wiker H.G. Comparison of membrane proteins of *Mycobacterium tuberculosis* H37Rv and H37Ra strains. *BMC Microbiol.*, 2011, vol. 11, no. 18, pp. 1-10.

12. Verma R., Pinto S.M., Patil A.H., et al. Quantitative Proteomic and Phosphoproteomic Analysis of H37Ra and H37Rv Strains of *Mycobacterium tuberculosis*. *J Proteome Res.*, 2017, vol. 16, no. 4, pp. 1632-1645. doi:10.1021/acs.jproteome.6b00983

13. Mattow J., Schaible U.E., Schmidt F. et al. Comparative proteome analysis of culture supernatant proteins from virulent *Mycobacterium tuberculosis* H37Rv and attenuated *M. bovis* BCG Copenhagen. *Electrophoresis*, 2003, vol. 24, no. 19-20, pp. 3405-3420. doi:10.1002/elps.200305601

14. Tan T., Lee W.L., Alexander D.C., Grinstein S., Liu J. *Cell Microbiol.,* 2006, vol. 8, no. 9, pp. 1417–1429.

15. Gunawardena H. P., Feltcher M. E., Wrobel J. A., Gu Sh., Braunstein M., Chen X. Comparison of the Membrane Proteome of Virulent *Mycobacterium tuberculosis* and the Attenuated *Mycobacterium bovis* BCG Vaccine Strain by Label-free Quantitative Proteomics. *J. Proteome Res.,* 2013, vol. 12, no.12, pp. 5463–5474. doi:10.1021/pr400334k.

16. Schubert O.T., Mouritsen J., Ludwig C. et al. The Mtb proteome library: A resource of assays to quantify the complete proteome of *Mycobacterium tuberculosis*. *Cell Host Microbe*., 2013, vol. 13, no. 5, pp. 602-612. doi:10.1016/j.chom.2013.04.008

17. SRM Atlas. Available at: http://www.srmatlas.org/ (accessed at 20 September 2018).

18. Picotti P., Bodenmiller B., Mueller L.N. et al. Full dynamic range proteome analysis of s. cerevisiae by targeted proteomics. *Cell*, 2009,vol. 138, pp. 795-806.

19. de Souza G.A., Wiker H.G. A proteomic view of mycobacteria. *Proteomics*, 2011, vol. 1, no. 15, pp. 3118-3127. doi:10.1002/pmic.201100043

20. Calder B., Albeldas C., Blackburn J.M., Soares N.C. Mass spectrometry offers insight into the role of ser/thr/tyr phosphorylation in the mycobacteria. *Front Microbiol*., 2016, vol. 7, no. 141, pp.1-8. doi:10.3389/fmicb.2016.00141

21. Nakedi K. C., Nel A. J., Garnett S., Blackburn J. M., Soares, N. C. Comparative Ser/Thr/Tyr phosphoproteomics between two mycobacterial species: the fast growing *Mycobacterium smegmatis* and the slow growing *Mycobacterium bovis* BCG. *Front. Microbiol.*, 2015, vol. 6, no. 237, pp.1-12 doi: 10.3389/fmicb.2015.00237

22. Prisic S., Dankwa S., Schwartz D. et al. Extensive phosphorylation with overlapping specificity by *Mycobacterium tuberculosis* serine/threonine protein kinases. *Proc Natl Acad Sci*., 2010, ;vol. 07, no. 16, pp. 7521-7526. doi:10.1073/pnas.0913482107

23. Kusebauch U., Ortega C., Ollodart A., Rogers R. S., Sherman D. R., Moritz R. L. et al. Mycobacterium tuberculosis supports protein tyro- sine phosphorylation. *Proc. Natl. Acad. Sci. U.S.A.* 2014, vol. 111, pp. 9265–9270. doi: 10.1073/pnas.1323894111

24. Ren J., Sang Y., Lu J., Yao Y.F. Protein Acetylation and Its Role in Bacterial Virulence. *Trends Microbiol*., 2017, vol. 25, no. 9, pp. 768-779. doi:10.1016/j.tim.2017.04.001

25. Okkels L.M., Müller E.C., Schmid M. et al. CFP10 discriminates between nonacetylated and acelylated ESAT-6 of *Mycobacterium tuberculosis* by differential interaction. *Proteomics*. 2004;4(10):2954-2960. doi:10.1002/pmic.200400906

26. Liu F., Yang M., Wang X. et al. Acetylome Analysis Reveals Diverse Functions of Lysine Acetylation in *Mycobacterium tuberculosis*. *Mol Cell Proteomics*, 2014, vol. 13, no. 12, pp. 3352-3366. doi:10.1074/mcp.M114.041962

27. Bi J., Wang Y., Yu H. et al. Modulation of Central Carbon Metabolism by Acetylation of Isocitrate Lyase in *Mycobacterium tuberculosis*. *Sci Rep*., 2017, vol. 7, pp. 1-11. doi:10.1038/srep44826

28. Yang H., Sha W., Liu Z. et al. Lysine acetylation of DosR regulates the hypoxia response of *Mycobacterium tuberculosis* article. *Emerg Microbes Infect*., 2018, vol. 7, no. 34, pp. 1-14. doi:10.1038/s41426-018-0032-2

29. Geoffrey T., Michael J., Sweredoski. O-linked glycosylation sites profiling in *Mycobacterium tuberculosis* culture filtrate proteins. *J. Proteomics*, 2014, vol. 97, pp. 296–306. doi:10.1016/j.jprot.2013.05.011.

30. González-Zamorano M., Hernández G.M., Xolalpa W. et al. *Mycobacterium tuberculosis* glycoproteomics based on ConA-lectin affinity capture of mannosylated proteins. *J Proteome Res*., 2009, vol. 8, no. 2, pp. 721-733. doi:10.1021/pr800756a

31. Liu C-F., Tonini L., Malaga W. et al. Bacterial protein-O-mannosylating enzyme is crucial for virulence of *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci.*, 2013, vol. 110, pp. 6560-6565.

32. Brudey K., Driscoll J. R., Rigouts L., Prodinger W. M., Gori A., Al-Hajoj S. A., Allix C. L.et al. *Mycobacterium tuberculosis* complex genetic diversity: mining the Fourth

International Spoligotyping Database (SpolDB4) for classification, population genetics and epidemiology. *BMC Microbiol.* 2006, vol. 6, no. 23, pp. 1-17.

33. Honisch C., Mosko M., Arnold C., Gharbia S.E., Diel R., Niemann S. Replacing reverse line blot hybridization spoligotyping of the *Mycobacterium tuberculosis* complex. *J Clin Microbiol.*, 2010, vol.48, no. 5, pp.1520-1526. doi:10.1128/JCM.02299-09

34. Shitikov E., Ilina E., Chernousova L. et al. Mass spectrometry based methods for the discrimination and typing of mycobacteria. *Infect. Genet. Evol.*, 2012, vol. 12, no. 4, pp. 838-845. doi:10.1016/j.meegid.2011.12.013